EXHIBIT I

OMNIBUS BROWN DECLARATION

Stata 13 help for regress

Title

[R] regress -- Linear regression

Syntax

regress *depvar* [*indepvars*] [*if*] [*in*] [*weight*] [, *options*]

options	Description
Model	
noconstant	suppress constant term
hascons	has user-supplied constant
tsscons	compute total sum of squares with constant; seldom used
SE/Robust	
vce(*vcetype***)**	*vcetype* may be **ols**, **robust**, **cluster** *clustvar*, **bootstrap**, **jackknife**, **hc2**, or **hc3**
Reporting	
level(#)	set confidence level; default is **level(95)**
beta	report standardized beta coefficients
eform(*string***)**	report exponentiated coefficients and label as *string*
depname(*varname***)**	substitute dependent variable name; programmer's option
display_options	control column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
noheader	suppress output header
notable	suppress coefficient table
plus	make table extendable
mse1	force mean squared error to **1**
coeflegend	display legend instead of statistics

indepvars may contain factor variables; see fvvarlist.
depvar and *indepvars* may contain time-series operators; see tsvarlist.
bootstrap, **by**, **fp**, **jackknife**, **mfp**, **mi estimate**, **nestreg**, **rolling**, **statsby**, **stepwise**, and **svy** are allowed; see prefix.
vce(bootstrap) and **vce(jackknife)** are not allowed with the **mi estimate** prefix.
Weights are not allowed with the **bootstrap** prefix.
aweights are not allowed with the **jackknife** prefix.
hascons, **tsscons**, **vce()**, **beta**, **noheader**, **notable**, **plus**, **depname()**, **mse1**, and weights are not allowed with the **svy** prefix.
aweights, **fweight**s, **iweight**s, and **pweight**s are allowed; see weight.
noheader, **notable**, **plus**, **mse1**, and **coeflegend** do not appear in the dialog

box.
See **[R] regress postestimation** for features available after estimation.

Menu

Statistics > Linear models and related > Linear regression

Description

regress fits a model of *depvar* on *indepvars* using linear regression.

Here is a short list of other regression commands that may be of interest. See estimation commands for a complete list.

Command	Description
areg	an easier way to fit regressions with many dummy variables
arch	regression models with ARCH errors
arima	ARIMA models
boxcox	Box-Cox regression models
cnsreg	constrained linear regression
eivreg	errors-in-variables regression
etregress	linear regression with endogenous treatment effects
frontier	stochastic frontier models
gmm	generalized method of moments estimation
heckman	Heckman selection model
intreg	interval regression
ivregress	single-equation instrumental-variables regression
ivtobit	tobit regression with endogenous variables
newey	regression with Newey-West standard errors
nl	nonlinear least-squares estimation
nlsur	estimation of nonlinear systems of equations
qreg	quantile (including median) regression
reg3	three-stage least-squares (3SLS) regression
rreg	a type of robust regression
gsem	generalized structural equation models
sem	linear structural equation models
sureg	seemingly unrelated regression
tobit	tobit regression
truncreg	truncated regression
xtabond	Arellano-Bond linear dynamic panel-data estimation
xtdpd	linear dynamic panel-data estimation
xtfrontier	panel-data stochastic frontier model
xtgls	panel-data GLS models
xthtaylor	Hausman-Taylor estimator for error-components models
xtintreg	panel-data interval regression models
xtivreg	panel-data instrumental-variables (2SLS) regression
xtpcse	linear regression with panel-corrected standard errors
xtreg	fixed- and random-effects linear models
xtregar	fixed- and random-effects linear models with an AR(1) disturbance
xttobit	panel-data tobit models

Options

+-------+
----+ Model +----------

noconstant; see **[R] estimation options**.

hascons indicates that a user-defined constant or its equivalent is specified among the independent variables in *indepvars*. Some caution is recommended when specifying this option, as resulting estimates may not be as accurate as they otherwise would be. Use of this option requires "sweeping" the constant last, so the moment matrix must be accumulated in absolute rather than deviation form. This option may be safely specified when the means of the dependent and independent variables are all reasonable and there is not much collinearity between the independent variables. The best procedure is to view **hascons** as a reporting option -- estimate with and without **hascons** and verify that the coefficients and standard errors of the variables not affected by the identity of the constant are unchanged.

tsscons forces the total sum of squares to be computed as though the model has a constant, that is, as deviations from the mean of the dependent variable. This is a rarely used option that has an effect only when specified with **noconstant**. It affects only the total sum of squares and all results derived from the total sum of squares.

+-----------+
----+ SE/Robust +----------

vce(*vcetype***)** specifies the type of standard error reported, which includes types that are derived from asymptotic theory (**ols**), that are robust to some kinds of misspecification (**robust**), that allow for intragroup correlation (**cluster** *clustvar*), and that use bootstrap or jackknife methods (**bootstrap**, **jackknife**); see **[R]** *vce_option*.

vce(ols), the default, uses the standard variance estimator for ordinary least-squares regression.

regress also allows the following:

vce(hc2) and **vce(hc3)** specify an alternative bias correction for the robust variance calculation. **vce(hc2)** and **vce(hc3)** may not be specified with the **svy** prefix. In the unclustered case, **vce(robust)** uses (sigma-hat_j)^2={n/(n-k)}(u_j)^2 as an estimate of the variance of the jth observation, where u_j is the calculated residual and n/(n-k) is included to improve the overall estimate's small-sample properties.

vce(hc2) instead uses u_j^2/(1-h_jj) as the observation's variance estimate, where h_jj is the diagonal element of the hat (projection) matrix. This estimate is unbiased if the model really is homoskedastic. **vce(hc2)** tends to produce slightly more conservative confidence intervals.

vce(hc3) uses u_j^2/(1-h_jj)^2 as suggested by Davidson and MacKinnon (1993), who report that this method tends to produce better results when the model really is heteroskedastic. **vce(hc3)** produces confidence intervals that tend to be even more conservative.

See Davidson and MacKinnon (1993, 554-556) and Angrist and Pischke (2009, 294-308) for more discussion on these two bias corrections.

+-----------+
----+ Reporting +----------

level(#); see **[R] estimation options**.

beta asks that standardized beta coefficients be reported instead of confidence intervals. The beta coefficients are the regression coefficients obtained by first standardizing all variables to have a mean of 0 and a standard deviation of 1. **beta** may not be specified with **vce(cluster** *clustvar***)** or the **svy** prefix.

eform(*string***)** is used only in programs and ado-files that use **regress** to fit models other than linear regression. **eform()** specifies that the coefficient table be displayed in exponentiated form as defined in **[R] maximize** and that *string* be used to label the exponentiated coefficients in the table.

depname(*varname***)** is used only in programs and ado-files that use **regress** to fit models other than linear regression. **depname()** may be specified only at estimation time. *varname* is recorded as the identity of the dependent variable, even though the estimates are calculated using *depvar*. This method affects the labeling of the output -- not the results calculated -- but could affect subsequent calculations made by **predict**, where the residual would be calculated as deviations from *varname* rather than *depvar*. **depname()** is most typically used when *depvar* is a temporary variable (see **[P] macro**) used as a proxy for *varname*.

depname() is not allowed with the **svy** prefix.

display_options: **noomitted**, **vsquish**, **noemptycells**, **baselevels**, **allbaselevels**, **nofvlabel**, **fvwrap(#)**, **fvwrapon(***style***)**, **cformat(***%fmt***)**, **pformat(%***fmt***)**, **sformat(%***fmt***)**, and **nolstretch**; see **[R] estimation options**.

The following options are available with **regress** but are not shown in the dialog box:

noheader suppresses the display of the ANOVA table and summary statistics at the top of the output; only the coefficient table is displayed. This option is often used in programs and ado-files.

notable suppresses display of the coefficient table.

plus specifies that the output table be made extendable. This option is often used in programs and ado-files.

mse1 is used only in programs and ado-files that use **regress** to fit models other than linear regression and is not allowed with the **svy** prefix. **mse1** sets the mean squared error to **1**, thus forcing the variance-covariance matrix of the estimators to be (X'DX)^-1 (see *Methods and formulas* in **[R] regress**) and affecting calculated standard errors. Degrees of freedom for t statistics is calculated as n rather than n-k.

coeflegend; see **[R] estimation options**.

Examples: linear regression

Setup
```
. sysuse auto
```

Fit a linear regression
```
. regress mpg weight foreign
```

Fit a better linear regression, from a physics standpoint
```
. gen gp100m = 100/mpg
. regress gp100m weight foreign
```

Obtain beta coefficients without refitting model
```
. regress, beta
```

Suppress intercept term
```
. regress weight length, noconstant
```

Model already has constant
```
. regress weight length bn.foreign, hascons
```

Examples: regression with robust standard errors

```
. sysuse auto, clear
. generate gpmw = ((1/mpg)/weight)*100*1000
. regress gpmw foreign
. regress gpmw foreign, vce(robust)
. regress gpmw foreign, vce(hc2)
. regress gpmw foreign, vce(hc3)
```

```
. webuse regsmpl, clear
. regress ln_wage age c.age#c.age tenure, vce(cluster id)
```

Example: weighted regression

```
. sysuse census
. regress death medage i.region [aw=pop]
```

Examples: linear regression with survey data

Setup
```
. webuse highschool
```

Perform linear regression using survey data
```
. svy: regress weight height
```

Setup
```
. generate male = sex == 1 if !missing(sex)
```

Perform linear regression using survey data for a subpopulation
```
. svy, subpop(male): regress weight height
```

Video example

Simple linear regression in Stata

Stored results

regress stores the following in **e()**:

Scalars

e(N)	number of observations
e(mss)	model sum of squares
e(df_m)	model degrees of freedom
e(rss)	residual sum of squares
e(df_r)	residual degrees of freedom
e(r2)	R-squared
e(r2_a)	adjusted R-squared
e(F)	F statistic
e(rmse)	root mean squared error
e(ll)	log likelihood under additional assumption of i.i.d. normal errors
e(ll_0)	log likelihood, constant-only model
e(N_clust)	number of clusters
e(rank)	rank of **e(V)**

Macros

e(cmd)	**regress**
e(cmdline)	command as typed
e(depvar)	name of dependent variable
e(model)	**ols** or **iv**
e(wtype)	weight type
e(wexp)	weight expression
e(title)	title in estimation output when **vce()** is not **ols**
e(clustvar)	name of cluster variable
e(vce)	*vcetype* specified in **vce()**
e(vcetype)	title used to label Std. Err.
e(properties)	**b V**
e(estat_cmd)	program used to implement **estat**
e(predict)	program used to implement **predict**
e(marginsok)	predictions allowed by **margins**
e(asbalanced)	factor variables **fvset** as **asbalanced**
e(asobserved)	factor variables **fvset** as **asobserved**

Matrices

e(b)	coefficient vector
e(V)	variance-covariance matrix of the estimators
e(V_modelbased)	model-based variance

Functions

e(sample)	marks estimation sample

References

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.